

CLAIMS

What is claimed is:

1. A method for retrieving a desired postal address from a plurality of postal addresses,
5 the method comprising:

a) receiving input data having one or more input terms for finding the desired postal address from a first database including data representing the plurality of postal addresses, each postal address being formed of one or more postal address elements;

b) searching a dictionary for entries in the dictionary corresponding to one or
10 more of the input terms, each entry in the dictionary corresponding to at least one of the postal address elements of the first database, wherein the dictionary is in the form of a tree data structure having a plurality of nodes including a root node and a multiplicity of leaf nodes, a path from the root node to one of the leaf nodes representing one of the postal address elements; and

c) outputting data representing a postal address, if any, represented in the first
15 database that has postal address elements that correspond with each of the entries of the dictionary that correspond to the input data.

2. The method of claim 1, wherein the first database is formed as a tree data structure having a plurality of nodes including a root node and a multiplicity of leaf nodes, a path from
20 the root node to one of the leaf nodes being representative of one of the postal addresses.

3. The method of claim 2, further comprising, before outputting data representing the postal address, determining whether a pair of nodes in the first database at different locations in the tree data structure relate to the same postal address.

4. The method of claim 1, wherein the data representing postal addresses includes, for
25 each postal address element, a code for the postal address element.

5. The method of claim 1, wherein the dictionary is arranged so that each particular node of the tree data structure under the root node contains a portion of at least one of the entries in the dictionary, the entry sharing a stem defined by a path from the root node to the particular node, and so that a plurality of the nodes have a plurality of the portions.

6. The method of claim 5, wherein each portion acts as one of a termination point and a
30 single path leading from that portion to another node.

7. The method of claim 5, wherein each portion of at least one node is a single character.

8. The method of claim 5, wherein at least one portion of at least one node includes a plurality of characters.

9. The method of claim 1, further comprising:

ascertaining a location for each occurrence of data within the first database

5 corresponding to at least one of the entries in the dictionary determined to correspond to one or more of the input terms; and

determining from the locations whether at least one of the postal addresses is in accordance with the input data.

10. The method of claim 1, wherein one of the entries in the dictionary corresponds to one or more of the input terms if the one or more input terms are identical to the entry.

11. The method of claim 1, wherein one of the entries in the dictionary corresponds to one or more of the input terms if the one or more input terms have a lower quality correspondence with the entry than being identical.

12. The method of claim 1, further comprising, if one or more entries in the dictionary are identical to at least one, but not all, of the input terms, thereby leaving one or more unmatched input terms, searching the dictionary for entries having a lower quality correspondence with the unmatched input terms.

13. The method of claim 1, wherein the postal address elements forming the postal addresses in the first database are divided into categories, and wherein data associated with a given postal address element indicates the category of the given postal address element.

14. The method of claim 13, wherein the categories are selected from the group consisting of a name element, a house number element, a street name element, a town element, a county element, a state element, and a postal code element.

15. The method of claim 13, wherein the data associated with the given postal address element that indicates the category is provided by the structure of the first database.

16. The method of claim 13, wherein the data associated with the given postal address element is provided by a number of characters available to represent the given postal address element.

17. The method of claim 13, wherein receiving input data having one or more input terms includes receiving an indication of the category of postal address element that each of the one or more input terms represents.

18. The method of claim 1, further comprising preprocessing at least one of the input terms to increase the likelihood of locating the desired postal address before searching the dictionary.

19. The method of claim 18, wherein preprocessing includes accounting for differences in data syntax between the input terms and the postal address elements represented by the data of at least one of the first database and the dictionary.

20. The method of claim 18, wherein preprocessing includes ascertaining whether one or more of the input terms correspond to a category of postal address element.

21. The method of claim 20, further comprising, if any of the input terms corresponds to one of the categories, including an indication of the category in the input data.

22. The method of claim 1, wherein the input data is entered manually by a user.

23. The method of claim 1, wherein the input data is received from a separate data store.

24. An apparatus for retrieving a desired postal address from a plurality of postal addresses, the apparatus comprising:

a memory storing a program;

a first database including data representing the plurality of postal addresses, each postal address being formed of one or more postal address elements;

a dictionary of entries, each entry in the dictionary corresponding to at least one of the postal address elements of the first database, wherein the dictionary is in the form of a tree data structure having a plurality of nodes including a root node and a multiplicity of leaf nodes, a path from the root node to one of the leaf nodes representing one of the postal address elements;

a processor in communication with the memory; in which the processor is directed by the program to:

receive input data having one or more input terms for finding the desired postal address from the first database;

search the dictionary for entries in the dictionary corresponding to one or more of the input terms; and

output data representing a postal address, if any, represented in the first database that has postal address elements that correspond with each of the entries of the dictionary that correspond to the input data.

25. A computer-readable medium for retrieving a desired postal address from a plurality of postal addresses, the computer-readable medium storing instructions that direct a microprocessor to:

a) receive input data having one or more input terms for finding the desired postal address from a first database including data representing the plurality of postal addresses, each postal address being formed of one or more postal address elements;

b) search a dictionary for entries in the dictionary corresponding to one or more of the input terms, each entry in the dictionary corresponding to at least one of the postal address elements of the first database, wherein the dictionary is in the form of a tree data structure having a plurality of nodes including a root node and a multiplicity of leaf nodes, a path from the root node to one of the leaf nodes representing one of the postal address elements; and

c) output data representing a postal address, if any, represented in the first database that has postal address elements that correspond with each of the entries of the dictionary that correspond to the input data.

26. A computer-readable medium that stores data accessible by a program executable on a data processing system, the data being organised according to a data structure that includes:

data in a first database representing a plurality of postal addresses, each postal address being formed of one or more postal address elements;

a dictionary of entries, each entry in the dictionary corresponding to at least one of the postal address elements of the first database, wherein the dictionary is in the form of a tree data structure having a plurality of nodes including a root node and a multiplicity of leaf nodes, a path from the root node to one of the leaf nodes representing one of the postal address elements; and

data linking a given entry in the dictionary with one or more of the postal addresses in the first database having a postal address element that corresponds to the given entry in the dictionary.

27. An apparatus for retrieving a desired postal address from a plurality of postal addresses, the apparatus comprising:

a) means for receiving input data having one or more input terms for finding the desired postal address from a first database including data representing the plurality of postal addresses, each postal address being formed of one or more postal address elements;

b) means for searching a dictionary for entries in the dictionary corresponding to one or more of the input terms, each entry in the dictionary corresponding to at least one of the postal address elements of the first database, wherein the dictionary is in the form of a tree data structure having a plurality of nodes including a root node and a multiplicity of leaf nodes, a path from the root node to one of the leaf nodes representing one of the postal address elements; and

c) means for outputting data representing a postal address, if any, represented in the first database that has postal address elements that correspond with each of the entries of the dictionary that correspond to the input data.

28. A method for retrieving a desired postal address from a plurality of postal addresses, the method comprising:

a) receiving input data having one or more input terms for finding the desired postal address from a first database including the plurality of postal addresses, each postal address being formed of one or more postal address elements;

b) searching a dictionary for entries in the dictionary corresponding to the one or more input terms, each entry in the dictionary corresponding to at least one of the postal address elements of the first database, wherein one of the entries in the dictionary corresponds to one of the input terms if the input term is identical to the entry or has a lower quality correspondence with the entry; and

c) outputting data representing the postal address, if any, represented in the first database that has postal address elements that correspond with each of the entries of the dictionary that corresponds to one or more of the input terms.

29. The method of claim 28, wherein searching the dictionary for entries includes, if one or more entries in the dictionary are identical to at least one, but not all, of the input terms, thereby leaving one or more unmatched input terms, searching the dictionary for entries having a lower quality correspondence with the unmatched input terms.

30. The method of claim 28, wherein the dictionary is in the form of a tree data structure having a plurality of nodes including a root node and a multiplicity of leaf nodes, a path from the root node to one of the leaf nodes representing one of the postal address elements.

31. The method of claim 30, wherein the dictionary is arranged so that each particular node of the tree data structure under the root node contains a portion of at least one of the

entries in the dictionary, the entry sharing a stem defined by a path from the root node to the particular node, and so that a plurality of the nodes have a plurality of the portions.

32. The method of claim 31, wherein each portion acts as one of a termination point and a single path leading from that portion to another node.

5 33. The method of claim 28, further comprising preprocessing at least one of the input terms to increase the likelihood of locating the desired postal address before searching the dictionary.

34. The method of claim 33, wherein preprocessing includes accounting for differences in data syntax between the input terms and the postal address elements represented by the data of
10 at least one of the first database and the dictionary.

35. An apparatus for retrieving a desired postal address from a plurality of postal addresses, the apparatus comprising:

a memory storing a program;

15 a first database including data representing the plurality of postal addresses, each postal address being formed of one or more postal address elements;

a dictionary of entries, each entry in the dictionary corresponding to at least one of the postal address elements of the first database;

a processor in communication with the memory; in which the processor is directed by the program to:

20 receive input data having one or more input terms for finding the desired postal address from the first database;

search the dictionary for entries in the dictionary corresponding to the one or more input terms, wherein one of the entries in the dictionary corresponds to one of the input terms if the input term is identical to the entry or has a lower quality correspondence with the
25 entry; and

output data representing the postal address, if any, represented in the first database that has postal address elements that correspond with each of the entries of the dictionary that corresponds to one or more of the input terms.

36. A computer-readable medium for retrieving a desired postal address from a plurality of
30 postal addresses, the computer-readable medium storing instructions that direct a microprocessor to:

a) receive input data having one or more input terms for finding the desired postal address from a first database including the plurality of postal addresses, each postal address being formed of one or more postal address elements;

b) search a dictionary for entries in the dictionary corresponding to the one or more input terms, each entry in the dictionary corresponding to at least one of the postal address elements of the first database, wherein one of the entries in the dictionary corresponds to one of the input terms if the input term is identical to the entry or has a lower quality correspondence with the entry; and

c) output data representing the postal address, if any, represented in the first database that has postal address elements that correspond with each of the entries of the dictionary that corresponds to one or more of the input terms.

37. An apparatus for retrieving a desired postal address from a plurality of postal addresses, the apparatus comprising:

a) means for receiving input data having one or more input terms for finding the desired postal address from a first database including the plurality of postal addresses, each postal address being formed of one or more postal address elements;

b) means for searching a dictionary for entries in the dictionary corresponding to the one or more input terms, each entry in the dictionary corresponding to at least one of the postal address elements of the first database, wherein one of the entries in the dictionary corresponds to one of the input terms if the input term is identical to the entry or has a lower quality correspondence with the entry; and

c) means for outputting data representing the postal address, if any, represented in the first database that has postal address elements that correspond with each of the entries of the dictionary that corresponds to one or more of the input terms.

38. A method for retrieving a desired postal address from a plurality of postal addresses, the method comprising:

a) receiving input data having one or more input terms for finding the desired postal address from a first database including the plurality of postal addresses, each postal address being formed of one or more postal address elements, wherein the first database is formed as a tree data structure having a plurality of nodes including a root node and a multiplicity of leaf nodes, a path from the root node to one of the leaf nodes being representative of one of the postal addresses;

b) searching a dictionary for entries in the dictionary corresponding to the one or more input terms, each entry in the dictionary corresponding to at least one of the postal address elements of the first database;

c) matching each entry in the dictionary that corresponds to one or more of the input terms with one or more postal address elements of one or more postal addresses in the first database;

d) determining, with use of the organization of the nodes in the first database, that a particular postal address in the first database has postal address elements that match with each of the entries of the dictionary that corresponds to one or more of the input terms; and

e) outputting data representing the particular postal address.

39. The method of claim 38, wherein the nodes of the first database are organized in a linear store.

40. The method of claim 39, wherein all of the descendants of any given node are located after that given node in the linear store, and before the next node in the linear store that is not a descendent of the given node.

41. The method of claim 40, wherein determining that a particular postal address in the first database has postal address elements that match with each of the entries of the dictionary that corresponds to one or more of the input terms includes assessing whether nodes, representing respective postal address elements, at different locations in the first database relate to the same postal address by using information concerning the respective locations in the linear store of the nodes.

42. The method of claim 41, wherein there is associated, with each node in the first database, data representative of a preset distance such that the determination of whether nodes in the first database relate to the same postal address can be performed by comparing the locations of the nodes with at least one of the respective preset distances associated with each of the nodes.

43. An apparatus for retrieving a desired postal address from a plurality of postal addresses, the apparatus comprising:

a memory storing a program;

a first database including the plurality of postal addresses, each postal address being formed of one or more postal address elements, wherein the first database is formed as a tree data structure having a plurality of nodes including a root node and a multiplicity of leaf

nodes, a path from the root node to one of the leaf nodes being representative of one of the postal addresses;

a dictionary of entries, each entry in the dictionary corresponding to at least one of the postal address elements of the first database;

5 a processor in communication with the memory; in which the processor is directed by the program to:

receive input data having one or more input terms for finding the desired postal address from the first database;

10 search the dictionary for entries in the dictionary corresponding to the one or more input terms;

match each entry in the dictionary that corresponds to one or more of the input terms with one or more postal address elements of one or more postal addresses in the first database;

15 determine, with use of the organization of the nodes in the first database, that a particular postal address in the first database has postal address elements that match with each of the entries of the dictionary that corresponds to one or more of the input terms; and

output data representing the particular postal address.

44. A computer-readable medium for retrieving a desired postal address from a plurality of postal addresses, the computer-readable medium storing instructions that direct a

20 microprocessor to:

a) receive input data having one or more input terms for finding the desired postal address from a first database including the plurality of postal addresses, each postal address being formed of one or more postal address elements, wherein the first database is formed as a tree data structure having a plurality of nodes including a root node and a multiplicity of leaf nodes, a path from the root node to one of the leaf nodes being representative of one of the postal addresses;

b) search a dictionary for entries in the dictionary corresponding to the one or more input terms, each entry in the dictionary corresponding to at least one of the postal address elements of the first database;

30 c) match each entry in the dictionary that corresponds to one or more of the input terms with one or more postal address elements of one or more postal addresses in the first database;

d) determine, with use of the organization of the nodes in the first database, that a particular postal address in the first database has postal address elements that match with each of the entries of the dictionary that corresponds to one or more of the input terms; and

e) output data representing the particular postal address.

5 45. A computer readable medium that stores data accessible by a program executable on a data processing system, the data being organized according to a data structure that includes:

a) data in a first database representing a plurality of postal addresses, each postal address being formed of one or more postal address elements, wherein the first database is formed as a tree data structure having a plurality of nodes including a root node and a
10 multiplicity of leaf nodes, a path from the root node to one of the leaf nodes being representative of one of the postal addresses, the nodes being organized in a linear store such that all of the descendants of any given node are located after that given node in the linear store, and before the next node in the linear store that is not a descendent of the given node;

b) a dictionary of entries, each entry in the dictionary corresponding to at least one
15 of the postal address elements of the first database; and

c) data associated with the entries in the dictionary, the data providing for each given entry in the dictionary: i) the one or more locations of the nodes in the first database corresponding to the postal address element that corresponds to the given entry in the dictionary; and ii) in respect of each location, an off set value representing the distance to the
20 next node, if any, in the linear store that is not a descendant of the node at that location, whereby the organization of the first database is such that said data may be used to assess whether a pair of nodes of the first database, at different locations in the linear store relate to the same postal address.

46. An apparatus for retrieving a desired postal address from a plurality of postal
25 addresses, the apparatus comprising:

a) means for receiving input data having one or more input terms for finding the desired postal address from a first database including the plurality of postal addresses, each postal address being formed of one or more postal address elements, wherein the first database is formed as a tree data structure having a plurality of nodes including a root node and a
30 multiplicity of leaf nodes, a path from the root node to one of the leaf nodes being representative of one of the postal addresses;

b) means for searching a dictionary for entries in the dictionary corresponding to the one or more input terms, each entry in the dictionary corresponding to at least one of the postal address elements of the first database;

c) means for matching each entry in the dictionary that corresponds to one or more of the input terms with one or more postal address elements of one or more postal addresses in the first database;

d) means for determining, with use of the organization of the nodes in the first database, that a particular postal address in the first database has postal address elements that match with each of the entries of the dictionary that corresponds to one or more of the input
10 terms; and

e) means for outputting data representing the particular postal address.

THE BOSTON HERALD